

# ARE QUANTUM CORRELATIONS GENUINELY QUANTUM?

Antonio Di Lorenzo

Instituto de Física, Universidade Federal de Uberlândia,  
Av. João Naves de Ávila 2121, Uberlândia, Minas Gerais, 38400-902, Brazil  
dilorenzo@infis.ufu.br

## Abstract

It is shown that the probabilities for the spin singlet can be reproduced through classical resources, with no communication between the distant parties, by using merely shared (pseudo-)randomness. If the parties are conscious beings aware of both the hidden-variables and the random mechanism, then one has a conspiracy. If the parties are aware of only the random variables, they may be induced to believe that they are able to send instantaneous information to one another. It is also possible to reproduce the correlations at the price of reducing the detection efficiency. It is further demonstrated that the same probability decomposition could be realized through action-at-a-distance, provided it existed.

## 1 Introduction and Outlook

Entangled states, first considered in Ref. 1, show one of the most challenging feature of quantum mechanics: unmeasured observables have no value. This fact manifests in peculiar probability distributions that do not admit embedding in a master, positive-definite probability that assigns simultaneous values to incompatible observables. While this statement applies as well to single particle states, that do not have in general a positive Wigner function, this feature becomes particularly interesting when components of an entangled state are separated in space. Then it is possible to establish the incompatibility between quantum mechanics and not one, but whole families of possible more fundamental theories, called hidden variable theories, that satisfy some general hypotheses. This remarkable feat was achieved by Bell,<sup>2</sup> who showed that a family of deterministic theories predicts that spin correlators satisfy an inequality, while the corresponding correlators predicted by quantum mechanical do not. Later the result was extended to a family of stochastic theories.<sup>3</sup> Experimental evidence excludes these theories,<sup>4–10</sup> so that, if a completion of quantum mechanics is to be found, one must look at theories that violate at least one of the hypotheses leading to Bell inequality. While locality, intended here as the impossibility

of action-at-a-distance, is one of these hypotheses, it is not the only one, despite a widespread belief that the present manuscript will prove unfounded. We are left then with the challenge of finding theories that reproduce the quantum mechanical correlations. In this paper, we concentrate on theories that violate the hypothesis of Uncorrelated Choice, which establishes the absence of correlations between the settings of the measuring instrument and the hidden variables. This violation admits several interpretations: a conspiracy between the measuring parties, a limitation of their free will, a non-local influence of the settings on the hidden variable, and also a conspiracy between the particles. We discuss also how to verify which interpretation is the correct one. In particular, we propose an (ideal) experiment allowing to decide whether we are in presence of non-locality or of limited free will, assuming that the hidden variables are experimentally accessible (if they were not, why bother?). This experiment, however, cannot rule out a superdeterminist theory, where a cosmic choreographer has decided from the beginning of times not only how experimentalists will choose the settings of their apparatuses (or which apparently-to-them random devices they select), but also all arbitrary conventions assigning meaning to, e.g., spin up and spin down, etc.\* Finally, given that the model we propose can be realized through local classical resources, and that it reproduces the quantum mechanical probabilities, it is to be concluded that quantum correlations are not exclusively quantum.

After reviewing the typical setup in Section 2, in Section 3 we discuss the hypotheses at the basis of the derivation of Bell and Leggett-type inequalities, dispelling some misconceptions about their physical meaning. Section 4 presents a simpler derivation of Bell inequality, based on a procedure pioneered by Boole relying only on the hypothesis of counter-factual definiteness. This derivation actually predates Bell's, but it was discovered in a mathematical context, where it was not realized the depth of its implications in physics, until it was rediscovered several years later. Furthermore, we discuss how this alternative derivation has a great relevance in refuting the Einstein-Podolski-Rosen argument, as it gives quantitative support to the criticisms of Bohr and others. Section 5 analyzes the Toner and Bacon model,<sup>12</sup> showing that it can be obtained as the limiting case of two different families of models, each violating a different Bell hypotheses, and also that it can be realized in a natural way through a limitation of the possible choices of polarization at one station. Section 6 presents a model by M.J.W. Hall,<sup>13</sup> based on minimal correlations between the hidden parameters and the choice of polarization. Section 7 introduces a model<sup>14</sup> violating only one of the Bell hypotheses and retaining Malus's law. It is also discussed how the model proposed can be explained in various ways — which include “conspiracy”, limitation of “free will”, exploitation of the detection loophole, or

---

\*Superdeterminism is unfalsifiable, but I assign it a very small probability of being true. It is akin to solipsism or intelligent design, in that it can explain anything but predict nothing. With Schopenhauer (who was talking about solipsism), I would say that superdeterminism should be treated as a “small frontier fortress. Admittedly the fortress is impregnable, but the garrison can never sally forth from it, and therefore we can pass it by and leave it in our rear without danger.”<sup>11</sup>

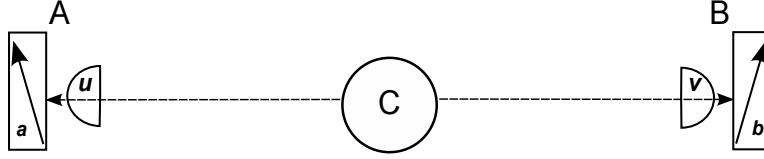


Figure 1: Scheme of the setup considered. Regions A and B are spacelike separated. The semicircles represent hypothetical detectors for the hidden variables as discussed in the Sec. 9.

action-at-a-distance,— depending on what additional hypotheses are made on the physical nature of the hidden variables. The two models of Section 6 and Section 7 are combined in Section 8, yielding a model that can violate Cirel’son bound and reach the maximum violation of Bell inequality. Section 9 discusses the relation between non-locality and “slave will”, and shows that it is possible to distinguish between the two, provided the hidden parameters can be measured, barring superdeterminism. Appendix 10 discusses competing definitions of locality, and shows with one example that one of them, which is used routinely in discussions of Bell inequality and that Bell argued coincided with “local causality”, conflicts with the other ones. Appendix 10 shows how it is possible to implement the models of Sections 6 and 7 through local classical resources, with no communication nor shared randomness between the distant parties, by only having shared randomness between each party and the entangler.

## 2 The system

The setup considered (see Fig. 1) consists of two particles produced in a region  $C$  of the spacetime, each travelling to a different detection region,  $A$  and  $B$ . The measurement of each particle does not need to happen at the same time, but the two detection events are assumed to have spacelike separation, so that a reference frame exists in which the measurements are simultaneous. The outcomes of the two measurements are two-valued, and will be denoted by  $e_A = \sigma = \pm 1$  and  $e_B = \tau = \pm 1$ . The  $A$  ( $B$ ) detector is characterized by a unit vector  $\mathbf{n}_A = \mathbf{a}$  ( $\mathbf{n}_B = \mathbf{b}$ ), corresponding to the orientation of a spin measuring device in Quantum Mechanics. The quantity of interest is the joint probability  $P(\sigma, \tau | \Psi, \Sigma)$  with  $\Psi$  describing the preparation of a singlet state, and  $\Sigma = \{\mathbf{a}, \mathbf{b}\}$  specifying the observables being measured. Since  $\Psi$  appears as a prior in all the probabilities considered, it is omitted for brevity. We shall write for brevity  $P(\sigma, \tau | \Psi, \Sigma) = P_{\sigma, \tau}(\Sigma)$ . A hidden variable theory consists in assuming the existence of additional parameters,  $\lambda$ , providing a finer description of the state of the system, so that, according to Bayes’s rule

$$P_{\sigma, \tau}(\Sigma) = \int d\mu(\lambda | \Sigma) P_{\sigma, \tau}(\lambda, \Sigma), \quad (1)$$

where  $\mu(\lambda|\Sigma) = \mu_\lambda(\Sigma)$  is the probability distribution for the additional parameters under the experimental conditions  $\Sigma = \{\mathbf{a}, \mathbf{b}\}$ . The joint probability  $P_{\sigma,\tau}(\lambda, \Sigma)$  can be further decomposed in either of two ways:

$$P_{\sigma,\tau}(\lambda, \mathbf{a}, \mathbf{b}) = M_\sigma^A(\lambda, \mathbf{a}, \mathbf{b})Q_\tau^B(\lambda, \mathbf{b}, \mathbf{a}, \sigma) = M_\tau^B(\lambda, \mathbf{b}, \mathbf{a})Q_\sigma^A(\lambda, \mathbf{a}, \mathbf{b}, \tau), \quad (2)$$

with the marginal probabilities

$$M_\sigma^A(\lambda, \mathbf{a}, \mathbf{b}) \equiv \sum_\tau P_{\sigma,\tau}(\lambda, \mathbf{a}, \mathbf{b}), \quad (3a)$$

$$M_\tau^B(\lambda, \mathbf{b}, \mathbf{a}) \equiv \sum_\sigma P_{\sigma,\tau}(\lambda, \mathbf{a}, \mathbf{b}), \quad (3b)$$

and the conditional probabilities

$$Q_\sigma^A(\lambda, \mathbf{a}, \mathbf{b}, \tau) \equiv \frac{P_{\sigma,\tau}(\lambda, \mathbf{a}, \mathbf{b})}{M_\tau^B(\lambda, \mathbf{b}, \mathbf{a})}, \quad (4a)$$

$$Q_\tau^B(\lambda, \mathbf{b}, \mathbf{a}, \sigma) \equiv \frac{P_{\sigma,\tau}(\lambda, \mathbf{a}, \mathbf{b})}{M_\sigma^A(\lambda, \mathbf{a}, \mathbf{b})}. \quad (4b)$$

For models respecting the symmetry between  $A$  and  $B$ , if  $B$  measured along  $\mathbf{a}$  and  $A$  along  $\mathbf{b}$  the new probabilities should coincide with the former ones, i.e.,  $M_\sigma^A(\lambda, \mathbf{a}, \mathbf{b}) = M_\sigma^B(\lambda, \mathbf{a}, \mathbf{b})$  and  $Q_\sigma^A(\lambda, \mathbf{a}, \mathbf{b}, \tau) = Q_\sigma^B(\lambda, \mathbf{a}, \mathbf{b}, \tau)$ , so that the upper index can be dropped since the functions coincide. There are models, however, where this symmetry is lost at the hidden variables level, see, e.g., the Toner and Bacon model discussed below. In the following, we shall write the hidden variables as  $\lambda = \lambda_A \cup \lambda_B \cup \lambda_G$ , where  $\lambda_j$  ( $j \in \{A, B\}$ ) refers to the local parameters, i.e. to the variables that can be associated, at the time of detection, to the regions  $A$  and  $B$  (they can be parameters locally attached to the particles and parameters describing the detection regions) and  $\lambda_G$  refers to the global parameters, i.e., parameters that it is not possible to pin down to a region of space-time, as the wavefunction of an entangled system. Bell<sup>15</sup> used the term “beables” to refer to local parameters, and assumed that there are no global parameters, presumably motivated by classical mechanics, where the local parameters are dynamical quantities describing physical systems, and no global parameters are needed if all the local ones are known with infinite precision.

### 3 Hypotheses at the basis of Bell and Leggett inequalities

#### 3.1 Bell inequality

Events corresponding to the detection of one particle, e.g., the determination of the spin projection of an electron along a given direction, can be reproduced easily by a hidden-variable model satisfying some reasonable requirements.<sup>2</sup>

However, when one considers a system composed of two entangled particles, reproducing the predictions of quantum mechanics becomes a non-trivial matter. Refs. 2 and 16 demonstrated that a family of deterministic and local models is incompatible with quantum mechanics, and the experiments<sup>4–10</sup> have corroborated this statement by reproducing the quantum mechanical predictions with remarkable accuracy. Refs. 2 and 16 discriminated between quantum mechanics and the said special class of deterministic local hidden-variable theories through inequalities, known as Bell and CHSH inequalities. The CHSH inequality, in particular, proves to be easier to check against experimental data, and it is currently written in a different form than in the original paper<sup>16</sup> as

$$|C(\mathbf{a}, \mathbf{b}) + C(\mathbf{a}', \mathbf{b}) + C(\mathbf{a}, \mathbf{b}') - C(\mathbf{a}', \mathbf{b}')| \leq 2, \quad (5)$$

with  $C$  the correlator (assuming  $\bar{\sigma} = \bar{\tau} = 0$ )

$$C(\mathbf{a}, \mathbf{b}) = \sum_{\sigma, \tau} \sigma \tau P_{\sigma, \tau}(\mathbf{a}, \mathbf{b}). \quad (6)$$

Under the hypothesis of determinism, the marginal probabilities are  $M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}) \in \{0, 1\}$  and  $M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}) \in \{0, 1\}$ . Then there is a discontinuous function  $S$  (respectively,  $T$ ) such that  $\sigma_0 = S(\lambda, \mathbf{a}, \mathbf{b})$  (respectively,  $\tau_0 = T(\lambda, \mathbf{b}, \mathbf{a})$ ) is the only permitted value of  $\sigma$  (resp.,  $\tau$ ). Hence, the conditional probabilities are

$$Q_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}, \tau_0) = Q_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}, T(\lambda, \mathbf{b}, \mathbf{a})) = M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}), \quad (7a)$$

$$Q_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}, \sigma_0) = Q_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}, S(\lambda, \mathbf{a}, \mathbf{b})) = M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}). \quad (7b)$$

Since the marginals are

$$M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}) = \sum_{\tau} P_{\sigma, \tau}(\lambda, \mathbf{a}, \mathbf{b}) = P_{\sigma, \tau_0}(\lambda, \mathbf{a}, \mathbf{b}) = M_{\tau_0}^B(\lambda, \mathbf{b}, \mathbf{a}) Q_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}, \tau_0) \quad (8a)$$

$$M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}) = \sum_{\sigma} P_{\sigma, \tau}(\lambda, \mathbf{a}, \mathbf{b}) = P_{\sigma_0, \tau}(\lambda, \mathbf{a}, \mathbf{b}) = M_{\sigma_0}^A(\lambda, \mathbf{a}, \mathbf{b}) Q_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}, \sigma_0), \quad (8b)$$

we have that for deterministic theories

$$Q_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}, \tau_0) = M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}), \quad (9a)$$

$$Q_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}, \sigma_0) = M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}). \quad (9b)$$

For  $\sigma \neq S(\lambda, \mathbf{a}, \mathbf{b})$  or  $\tau \neq T(\lambda, \mathbf{b}, \mathbf{a})$ , the conditional probabilities take the indeterminate form  $\frac{0}{0}$ . In any case, the joint probability, conditioned on  $\lambda$ , factorizes as

$$P_{\sigma, \tau}(\lambda, \mathbf{a}, \mathbf{b}) = M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}) M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}) \quad (10)$$

and the marginals are

$$M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}) = \delta_{\sigma, S(\lambda, \mathbf{a}, \mathbf{b})}, \quad M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}) = \delta_{\tau, T(\lambda, \mathbf{b}, \mathbf{a})} \quad (11)$$

Bell<sup>2</sup> then requires that the marginal probabilities obey the hypothesis of Setting Independence,

$$M_\sigma^A(\lambda, \mathbf{a}, \mathbf{b}) = M_\sigma^A(\lambda, \mathbf{a}) \quad (12a)$$

$$M_\tau^B(\lambda, \mathbf{b}, \mathbf{a}) = M_\tau^B(\lambda, \mathbf{b}) . \quad (12b)$$

For deterministic theories, Eqs. (12) are equivalent to  $S(\lambda, \mathbf{a}, \mathbf{b}) = S(\lambda, \mathbf{a})$ ,  $T(\lambda, \mathbf{b}, \mathbf{a}) = T(\lambda, \mathbf{b})$ . This hypothesis derives from the hypothesis of locality, meant as the impossibility of action-at-a-distance. Setting Independence is actually weaker than locality: let us imagine that the parameters  $\lambda$  can be divided into localized parameters associated to the particle reaching  $A$ ,  $\lambda_A$ , localized parameters associated to the particle reaching  $B$ ,  $\lambda_B$ , and global parameters  $\lambda_G$  that cannot be associated to a particle. Setting Independence allows the marginal probability at  $A$  to depend on the remote local parameters  $\lambda_B$ , even though it forbids dependence on the remote parameter  $\mathbf{b}$ .

Finally, it is required that

$$\mu(\lambda|\mathbf{a}, \mathbf{b}) = \mu(\lambda). \quad (13)$$

This hypothesis is sometimes called Measurement Independence, but we prefer to call it Uncorrelated Choice, meaning that the choice of the settings is uncorrelated to the value of  $\lambda$ . The original justification of Uncorrelated Choice relied on locality: if one imagines the  $\lambda$  as a deterministic time-evolution of  $\lambda_0$ , parameters localized at the entangler when the particles are emitted, changing the settings can not influence  $\lambda_0$  and hence it cannot influence  $\lambda$ . However, this picture was revealed<sup>17</sup> to be inexact, since a correlation between two events coinciding in space-time does not imply necessarily a causal relation between them, and even when there is such a causal relation, which event is the cause and which the effect is merely metaphysical (at least for time-symmetric physical theories such as classical and quantum mechanics). Indeed, by having  $\mu(\lambda|\mathbf{a}, \mathbf{b})$  we can think that the  $\lambda$  influence the choice of the settings, rather than *vice versa*. Locality implies not Eq. (13), but only that

$$\mu^A(\lambda_A|\lambda_G, \mathbf{a}, \mathbf{b}) = \frac{\int_{\Omega_B} d\lambda_B \mu(\lambda_A, \lambda_B, \lambda_G|\mathbf{a}, \mathbf{b})}{\int_{\Omega_A} d\lambda_A \int_{\Omega_B} d\lambda_B \mu(\lambda_A, \lambda_B, \lambda_G|\mathbf{a}, \mathbf{b})} = \mu^A(\lambda_A|\lambda_G, \mathbf{a}) \quad (14)$$

and

$$\mu^B(\lambda_B|\lambda_G, \mathbf{b}, \mathbf{a}) = \frac{\int_{\Omega_A} d\lambda_A \mu(\lambda_A, \lambda_B, \lambda_G|\mathbf{a}, \mathbf{b})}{\int_{\Omega_A} d\lambda_A \int_{\Omega_B} d\lambda_B \mu(\lambda_A, \lambda_B, \lambda_G|\mathbf{a}, \mathbf{b})} = \mu^B(\lambda_B|\lambda_G, \mathbf{b}), \quad (15)$$

where  $\Omega_j$  is the domain of  $\lambda_j$ . Thus the hypothesis of Uncorrelated Choice is stronger than locality.

In summary, the original derivation of Bell and CHSH inequalities relied on Determinism, which implies Eqs. (10) and (11), Setting Independence, which implies Eq. (12), and on Uncorrelated Choice, which requires Eq. (13). Since

Setting Independence and Uncorrelated Choice were believed to follow from locality, the inequalities, which are violated by quantum mechanics, were said to establish an incompatibility between local realistic theories and quantum mechanics. The term “realism” was used by Bell as synonymous with determinism. Later on, this statement was corrected to the incompatibility between quantum mechanics and all local deterministic theories allowing complete “free will”, i.e. satisfying Uncorrelated Choice.

While Eq. (11) was used in the original derivations of Refs. 2 and 16, later it was realized<sup>3,18</sup> that it was unnecessary, so that the probabilities were allowed to take values in the whole range  $[0, 1]$ . However, it was essential for deriving the inequalities, that Eq. (10), another consequence of determinism, still held. Confusingly, Eq. (10) was instead considered a consequence of locality, perhaps because it was implicitly assumed that the hidden variables consisted only of local parameters. It was then claimed that an incompatibility existed between all stochastic, local hidden-variable models and quantum mechanics. Ref. 19 pointed out, however, that Eq. (9) is not implied by locality, which, we remind the readers, means the impossibility of action-at-a-distance. Hence Eq. (9) or equivalently Eq. (10) has to be postulated for stochastic models. Ref. 19 called such hypothesis “completeness”, while Ref. 20 coined the name “Outcome Independence”. On the other hand, Bell<sup>15</sup> postulated “Outcome Independence” as the definition of “local causality”. Appendix 10 shows with an example that Bell’s definition is untenable, if one accepts that there are global parameters.

We shall refer to Eq. (10) as Reducibility of Correlations, since it means that the knowledge of  $\lambda$  breaks down the correlations between  $\sigma$  and  $\tau$ , so that correlations are reducible to ignorance of  $\lambda$ . Reducibility of Correlations implies that

$$Q_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}, \tau) = \begin{cases} M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}) , & \forall \tau : M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}) \neq 0 \\ \frac{0}{0} , & \forall \tau : M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}) = 0 \end{cases} , \quad (16a)$$

$$Q_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}, \sigma) = \begin{cases} M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}) , & \forall \sigma : M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}) \neq 0 \\ \frac{0}{0} , & \forall \sigma : M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}) = 0 \end{cases} . \quad (16b)$$

These equations characterize Outcome Independence, and they in turn imply Reducibility of Correlations. Hence the two terms are interchangeable.\*

In conclusion, Bell-type inequalities establish an incompatibility between quantum mechanics and a special class of stochastic models, the ones satisfying the hypotheses of Reducibility of Correlations, Setting Independence, and Uncorrelated Choice. While Ref. 19 and subsequent papers<sup>20–23</sup> highlighted the fact that Reducibility of Correlations is not a consequence of locality, these contributions were forgotten in the 90s, when the word “locality” in the context of Bell inequalities was redefined to mean compliance with the three hypotheses

---

\*Strictly speaking, if one interprets the symbol  $\frac{0}{0}$  as having a meaning of its own, then the term “Outcome Independence” does not describe properly Eqs. (16). However, one could as well say that since the symbol  $\frac{0}{0}$  is indeterminate, one may assign to it an arbitrary value,  $\frac{0}{0} = M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b})$  and  $\frac{0}{0} = M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a})$ .

Reducibility of Correlations, Setting Independence, and Uncorrelated Choice. This, I seem to understand, was due to the erroneous belief that the three hypotheses are implied by and imply the impossibility of action-at-a-distance. To make the situation worse, this special usage of “locality” has variances: some authors refer to Setting Independence as the no-signaling hypothesis (erroneously, since if  $B$  can change the local parameters  $\lambda_B$ , a message can be instantaneously sent to  $A$ ), and to Reducibility of Correlations as a consequence of “locality”; other authors refer to both Setting Independence and Reducibility of Correlations as “locality” conditions, others still take locality as synonymous with no-signaling. Throughout this paper we use locality as the impossibility of action-at-a-distance (see Appendix 10).

### 3.2 Leggett inequality

Ref. 24 considered a class of hidden-variable models which obey Uncorrelated Choice, as expressed by Eq. (13), and do not necessarily satisfy Reducibility of Correlations, but which obey an analogue of Malus’s law for the hidden variables, which consist in two unit vectors,  $\lambda = \{\mathbf{u}, \mathbf{v}\}$ :

$$M_\sigma^A(\lambda, \mathbf{a}, \mathbf{b}) = (1 + \sigma \mathbf{u} \cdot \mathbf{a})/2, \quad (17a)$$

$$M_\tau^B(\lambda, \mathbf{b}, \mathbf{a}) = (1 + \tau \mathbf{v} \cdot \mathbf{b})/2. \quad (17b)$$

It was shown<sup>24–26</sup> that these models predicts a correlator satisfying an inequality known as Leggett inequality, which is violated by quantum mechanics and by experiments.<sup>25, 27–32</sup> It is unfortunately confusing that the new Leggett inequality was presented as an incompatibility theorem between quantum mechanics and a class of non-local models. Indeed, models satisfying Eqs. (17) satisfy locality, if one thinks of  $\mathbf{u}$  and  $\mathbf{v}$  as parameters attached to each particle, while, as we discussed above, models satisfying Setting Independence in general do not. In this sense, Bell inequalities rule out also some non-local models, while Leggett inequality does not, contrary to what is claimed. The source of the misunderstanding is the following: in order to justify the correlation

$$P_{\sigma, \tau}(\lambda, \mathbf{a}, \mathbf{b}) \neq M_\sigma^A(\lambda, \mathbf{a}, \mathbf{b}) M_\tau^B(\lambda, \mathbf{b}, \mathbf{a}), \quad (18)$$

Ref. 24 assumed the existence of further variables  $\lambda'$ , in terms of which

$$P_{\sigma, \tau}(\lambda, \mathbf{a}, \mathbf{b}) = \int d\mu(\lambda'|\lambda) P_{\sigma, \tau}(\lambda, \lambda', \mathbf{a}, \mathbf{b}), \quad (19)$$

where the joint conditional probability is assumed to factorize as

$$P_{\sigma, \tau}(\lambda, \lambda', \mathbf{a}, \mathbf{b}) = \delta_{\sigma, S(\lambda, \lambda', \mathbf{a}, \mathbf{b})} \delta_{\tau, T(\lambda, \lambda', \mathbf{b}, \mathbf{a})}. \quad (20)$$

These models are non-local, since they violate Setting Independence, and deterministic. Leggett coined the term “crypto-non-local” theories to refer to them. Yet, it was realized that these models are only a special case of non-local models,<sup>33</sup> and that it is irrelevant that  $\lambda'$  exist or not, given that Eqs. (13) and (17)



imply Leggett inequality, independently of how they are arrived to.<sup>26,34</sup> Furthermore, one could say as well that the violation of Bell inequality rules out all the crypto-non-local theories such that upon integration of  $\lambda'$  Eqs. (10), (12), and (13) hold. Thus, the importance of Leggett inequality consists not in the fact that it discriminates between quantum mechanics and non-local theories, but in that it does not require Reducibility of Correlations, just Uncorrelated Choice and a more specific form of Setting Independence.

### 3.3 Models reproducing the quantum predictions

Examples that the violation of Uncorrelated Choice could lead to reproducing the quantum mechanical prediction were provided in Ref. 13,35. These models do not satisfy Malus's law. The amount of violation of Uncorrelated Choice necessary to reproduce quantum mechanics was recently quantified.<sup>13,36</sup> In Section 7, we discuss a model of ours<sup>14</sup> that satisfies at the same time the hypotheses of Malus's law (and hence Setting Independence) and Reducibility of Correlations, but violates Uncorrelated Choice. In the literature there are also models violating both Uncorrelated Choice and Reducibility of Correlations,<sup>25,37,38</sup> only Setting Independence,<sup>12,39</sup> or only Reducibility of Correlations.<sup>40</sup>

Clearly, by violating one of the hypotheses at the basis of Bell and Leggett inequalities it may be possible to violate them. Reproducing quantum mechanics, however, is not guaranteed, since the violation of Bell and Leggett inequalities is a necessary but not sufficient condition.

## 4 Alternative derivation of Bell inequalities

More than a hundred years before Bell derived the famous inequality, Boole<sup>41</sup> defined what he called 'conditions of possible experience': He imposed that the observed probabilities of  $k$  events  $E_{j_1}, E_{j_2}, \dots, E_{j_k}$  selected from a set  $\{E_1, \dots, E_n\}$  were the marginals of a master probability  $P(E_1, \dots, E_n)$ . This led to inequalities of the form

$$\sum_{r=1}^R A_r p_r \leq A_0 \quad (21)$$

where  $R$  is the number of different subsets of events for which there are experimental data, while  $p_r$  is the corresponding probability, and  $A_r$  are properly chosen coefficients. It is possible to derive Bell-type inequalities from Boole inequalities, as was done in various papers predating Bell's,<sup>42-45</sup> where Boole's inequalities were rediscovered independently. These papers, however, were concerned with mathematical and computational problems, and their authors were not aware of the implications of the results in physics. As an example, the

CHSH inequality can be derived in a few lines:

$$\begin{aligned}
& |C(\mathbf{a}, \mathbf{b}) + C(\mathbf{a}', \mathbf{b}) + C(\mathbf{a}, \mathbf{b}') - C(\mathbf{a}', \mathbf{b}')| \\
&= \left| \sum_{\substack{\sigma, \sigma' \\ \tau, \tau'}} (\sigma\tau + \sigma'\tau + \sigma\tau' - \sigma'\tau') P_{\sigma, \tau, \sigma', \tau'}(\mathbf{a}, \mathbf{b}, \mathbf{a}', \mathbf{b}') \right| \\
&\leq \sum_{\substack{\sigma, \sigma' \\ \tau, \tau'}} |\sigma\tau + \sigma'\tau + \sigma\tau' - \sigma'\tau'| P_{\sigma, \tau, \sigma', \tau'}(\mathbf{a}, \mathbf{b}, \mathbf{a}', \mathbf{b}') \\
&\leq \sum_{\substack{\sigma, \sigma' \\ \tau, \tau'}} (|\sigma + \sigma'| + |\sigma - \sigma'|) P_{\sigma, \tau, \sigma', \tau'}(\mathbf{a}, \mathbf{b}, \mathbf{a}', \mathbf{b}') = 2. \tag{22}
\end{aligned}$$

The beauty of this alternative route to Bell-type inequalities consists in that it does not rely on the existence of hidden variables and requires only one hypothesis, the ‘condition of possible experience’, or, in current terminology, counterfactual definiteness: spin components along any direction have a well defined probability distribution, independently of which component is being measured. This is but the definition of ‘elements of reality’ given in the paper by Einstein, Podolsky, and Rosen.<sup>1</sup> While it is dubious that the three hypotheses that Bell formulated in his derivation of the inequality and the very concept of hidden variables correspond to the hypotheses made in the EPR paper, counter-factual definiteness, instead, is certainly among the latter ones. Thus, the Boole-style’s derivation of the Bell-type inequalities, and the experimental violation of these, proves the EPR assumption of ‘elements of reality’ fallacious. This was indeed understood immediately, albeit in a qualitative way, in many responses<sup>46–51</sup> to the EPR paper, of which Bohr’s<sup>48</sup> is the most famous, but perhaps not the most clear.

I conclude this section acknowledging Refs. 52 and 53 that brought to the attention of the physics community the pioneering results of Boole and those of Bass, Schell, and Vorob’ev, respectively. Furthermore, the Boole-style derivation of the inequalities was rediscovered independently in Refs. 54 and 55. Finally, Ref. 56 proved that the observed frequencies of experiments (the frequency of observing, e.g., an outcome  $\sigma$  times the frequency with which the detector is chosen along a direction  $\mathbf{a}$ ) always admit a master probability distribution, i.e., they are Kolmogorovian in the terminology of Ref. 56). This shows quantitatively that superdeterminism can never be ruled out.

## 5 The Toner and Bacon model

In order to show an example of the interchangeability between non-locality and limited free will, let us consider the Toner and Bacon model,<sup>12</sup> which relies on a station sending a bit of information  $c$  to the other one, presumably by instantaneous communication, otherwise one could not violate Bell inequality in real time without keeping the locality loophole open. The model works in the

following way: both  $A$  and  $B$  can access two unit-vector hidden-variables  $\mathbf{u}, \mathbf{v}$ , which have a uniform distribution  $\mu_{\mathbf{u}, \mathbf{v}} = 1/(4\pi)^2$  independent of the settings. The outcome of  $A$  is deterministic, its marginal probability being

$$M_\sigma^A(\mathbf{a}, \mathbf{u}, \mathbf{v}) = \delta_{\sigma, S(\lambda, \mathbf{a})} \quad (23)$$

with  $S(\lambda, \mathbf{a}) = \text{sgn}(\mathbf{u} \cdot \mathbf{a})$ , and the sign function  $\text{sgn}(x) = -1, x < 0$  and  $\text{sgn}(x) = +1, x > 0$ . The conditional probability at  $B$  is

$$Q_\tau^B(\mathbf{b}, \mathbf{u}, \mathbf{v}, \mathbf{a}, \sigma) = \begin{cases} \frac{0}{0}, & \text{if } \sigma = -S(\lambda, \mathbf{a}) \\ \delta_{\tau, T(\lambda, \mathbf{a}, \mathbf{b})}, & \text{if } \sigma = S(\lambda, \mathbf{a}) \end{cases} \quad (24)$$

where the indeterminate form  $\frac{0}{0}$  reflects the fact that the priors  $\mathbf{u}, \mathbf{a}, \sigma$  are incompatible,  $T(\lambda, \mathbf{a}, \mathbf{b}) = -\text{sgn}[(\mathbf{u} + c(\lambda, \mathbf{a})\mathbf{v}) \cdot \mathbf{b}]$ ,  $c(\lambda, \mathbf{a}) = \text{sgn}(\mathbf{u} \cdot \mathbf{a}) \text{sgn}(\mathbf{v} \cdot \mathbf{a})$ .

### 5.1 Two possible extensions of the Toner and Bacon model.

Let us lift the hypothesis of determinism so that, instead of Eq. (23), we have

$$M_\sigma^A(\mathbf{a}, \mathbf{u}, \mathbf{v}) = p \delta_{\sigma, S(\lambda, \mathbf{a})} + (1-p) \delta_{\sigma, -S(\lambda, \mathbf{a})}, \quad (25)$$

where  $0 < p < 1$  is a fixed probability. There is an ambiguity about how to interpret the bit  $c$ . The latter, indeed, can be seen either as a function of  $\mathbf{u}, \mathbf{v}, \mathbf{a}$ , or, equivalently, as  $c = \tilde{c}(\mathbf{v}, \mathbf{a}, \sigma) = \sigma \text{sgn}(\mathbf{v} \cdot \mathbf{a})$ . Then Eq. (24) becomes either

$$Q_\tau^B(\mathbf{b}, \mathbf{u}, \mathbf{v}, \mathbf{a}, \sigma) = \delta_{\tau, T(\lambda, \mathbf{a}, \mathbf{b})}, \quad (26a)$$

if we choose  $c = c(\lambda, \mathbf{a})$ , or

$$Q_\tau^B(\mathbf{b}, \mathbf{u}, \mathbf{v}, \mathbf{a}, \sigma) = \delta_{\tau, \tilde{T}(\lambda, \mathbf{a}, \mathbf{b}, \sigma)}, \quad (26b)$$

with  $\tilde{T}(\lambda, \mathbf{a}, \mathbf{b}, \sigma) = -\text{sgn}[(\mathbf{u} + \tilde{c}(\mathbf{v}, \mathbf{a}, \sigma)\mathbf{v}) \cdot \mathbf{b}]$ , if we choose  $c = \tilde{c}(\mathbf{v}, \mathbf{a}, \sigma)$ . We have thus two families of hidden variable models, depending on a parameter  $p$ . In the first family, the conditional probability does not depend on  $\sigma$ , and it coincides with the marginal probability, which, hence, depends on  $\mathbf{a}$ , violating Setting Independence. After integrating out  $\mathbf{u}, \mathbf{v}$ , the probability is

$$P_{\sigma, \tau}(\mathbf{a}, \mathbf{b}) = \frac{1}{4} [1 - (2p-1)\sigma\tau \mathbf{a} \cdot \mathbf{b}]. \quad (27a)$$

In the second family, we have that Reducibility of Correlations is violated, instead, while the marginal probability at station  $A$  satisfies Setting Independence. Since the Toner and Bacon model has the (unaesthetic, in my opinion) feature of being asymmetric, if we instead consider the marginal probability at  $B$  and the conditional probability at  $A$ , we arrive to

$$M_\tau^B(\mathbf{b}, \mathbf{u}, \mathbf{v}, \mathbf{a}) = p \delta_{\tau, \tilde{T}(\lambda, \mathbf{a}, \mathbf{b}, S(\mathbf{u}, \mathbf{a}))} + (1-p) \delta_{\tau, \tilde{T}(\lambda, \mathbf{a}, \mathbf{b}, -S(\mathbf{u}, \mathbf{a}))},$$

and

$$Q_{\sigma}^A(\mathbf{a}, \mathbf{u}, \mathbf{v}, \mathbf{b}, \tau) = \frac{p \delta_{\sigma, S(\lambda, \mathbf{a})} \delta_{\tau, \tilde{T}(\lambda, \mathbf{a}, \mathbf{b}, S(\mathbf{u}, \mathbf{a}))} + (1-p) \delta_{\sigma, -S(\lambda, \mathbf{a})} \delta_{\tau, \tilde{T}(\lambda, \mathbf{a}, \mathbf{b}, -S(\mathbf{u}, \mathbf{a}))}}{M_{\tau}^B(\mathbf{b}, \mathbf{u}, \mathbf{v}, \mathbf{a})},$$

so that both Setting Independence and Reducibility of Correlations are violated by the second family in this alternative factorization. However, for asymmetric models, we propose to say that Setting Independence is violated when it is not satisfied in either of the two possible factorizations of the joint probability,  $P_{\sigma, \tau}(\lambda, \mathbf{a}, \mathbf{b}) = M_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}) Q_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a}, \sigma) = Q_{\sigma}^A(\lambda, \mathbf{a}, \mathbf{b}, \tau) M_{\tau}^B(\lambda, \mathbf{b}, \mathbf{a})$ . (If Reducibility of Correlations is satisfied in one factorization, it is also in the alternative one.) Therefore, in this second case, Setting Independence is satisfied. The resulting average probability is

$$P_{\sigma, \tau}(\mathbf{a}, \mathbf{b}) = \frac{1}{4} [1 - p\sigma\tau \mathbf{a} \cdot \mathbf{b}]. \quad (27b)$$

Both Eq. (27a) and (27b) reproduce the quantum mechanical probability for  $p \rightarrow 1$ , and both families of models tend to the Toner and Bacon model in this limit: the first family violates Setting Independence and preserves Reducibility of Correlations, the second family does the opposite. In a sense, models not complying with the Reducibility of Correlations are more robust, since they deviate from the quantum mechanical probability less than the models not complying with Setting Independence (correlations are reduced by  $p$  instead than by  $2p - 1$ ).

## 5.2 Local realization through limited free will

Finally, let us show that the Toner and Bacon model admits a simple reinterpretation in terms of limited free will. We assume that  $c$  is not a bit generated by  $A$ , but a binary hidden-variable, in addition to two unit vectors  $\mathbf{u}, \mathbf{v}$ , which are distributed uniformly.

When  $A$  receives  $\mathbf{u}, \mathbf{v}, c$ , the setting will be chosen so that  $c = c(\mathbf{u}, \mathbf{v}, \mathbf{a})$  holds. We may imagine that  $\sigma = S(\mathbf{u}, \mathbf{a})$  is enforced by some appropriate local physical law. Analogously,  $B$  chooses an orientation  $\mathbf{b}$ , but freely, while the physical laws on his side determine that  $\tau = T(\mathbf{u}, \mathbf{v}, c, \mathbf{b}) = S(-\mathbf{u} - c\mathbf{v}, \mathbf{b})$ . Hence, the hidden variables are, in this reformulation,  $\lambda = \{\mathbf{u}, \mathbf{v}, c\}$ , and they are distributed according to

$$\mu'_{\mathbf{u}, \mathbf{v}, c}(\mathbf{a}, \mathbf{b}) = \frac{1}{(4\pi)^2} \delta_{c, c(\mathbf{u}, \mathbf{v}, \mathbf{a})}. \quad (28)$$

The marginal probability at  $A$  is still

$$M_{\sigma}^A(\mathbf{u}, \mathbf{v}, c, \mathbf{a}) = \delta_{\sigma, S(\mathbf{u}, \mathbf{a})} \quad (29)$$

while the conditional probability at  $B$  is

$$Q_{\tau}^B(\mathbf{b}, \mathbf{u}, \mathbf{v}, \mathbf{a}, \sigma) = \begin{cases} \frac{0}{0}, & \text{if } \sigma = -S(\mathbf{u}, \mathbf{a}) \\ \delta_{\tau, T'(\lambda, \mathbf{b})}, & \text{if } \sigma = S(\mathbf{u}, \mathbf{a}) \end{cases} \quad (30)$$

with  $T'(\lambda, \mathbf{b}) = -\text{sgn}[(\mathbf{u} + c\mathbf{v}) \cdot \mathbf{b}]$ , so that both Setting Independence and Reducibility of Correlations are satisfied.

## 6 Hall model

Ref. 13 proposes a different decomposition of the quantum mechanical probability, through a model which has the maximum “free will” compatible with quantum mechanics, and deterministic outcomes:

$$\mu(\mathbf{u}, \mathbf{v} | \mathbf{a}, \mathbf{b}) = \delta(\mathbf{u} + \mathbf{v}) \frac{1 - f(\mathbf{u}, \mathbf{v}, \mathbf{a}, \mathbf{b})}{8 \arccos f(\mathbf{u}, \mathbf{v}, \mathbf{a}, \mathbf{b})}, \quad (31)$$

where

$$f(\mathbf{u}, \mathbf{v}, \mathbf{a}, \mathbf{b}) = \text{sgn}(\mathbf{u} \cdot \mathbf{a}) \text{sgn}(\mathbf{v} \cdot \mathbf{b}) \mathbf{a} \cdot \mathbf{b}, \quad (32)$$

and  $\text{sgn}(x) = 1$  for  $x \geq 0$ ,  $\text{sgn}(x) = -1$  for  $x < 0$ . The probability of the outcomes, given  $\mathbf{u}, \mathbf{v}$ , is

$$P_{\sigma, \tau}(\mathbf{u}, \mathbf{v}, \mathbf{a}, \mathbf{b}) = \delta_{\sigma, \text{sgn}(\mathbf{u} \cdot \mathbf{a})} \delta_{\tau, \text{sgn}(\mathbf{v} \cdot \mathbf{b})}, \quad (33)$$

so that both Setting Independence and Reducibility of Correlations are satisfied. We notice that if  $\mathbf{u}, \mathbf{v}$  are considered local parameters, localized, respectively, at the particle reaching  $A$  and the one reaching  $B$ , the model is non-local, because of the tangled nature of the function  $f$ . Indeed, the marginal distribution for  $\mathbf{u}$  is

$$\mu^A(\mathbf{u} | \mathbf{a}, \mathbf{b}) = \frac{1 - f(\mathbf{u}, -\mathbf{u}, \mathbf{a}, \mathbf{b})}{8 \arccos f(\mathbf{u}, -\mathbf{u}, \mathbf{a}, \mathbf{b})}, \quad (34)$$

and clearly depends on  $\mathbf{b}$ . Thus, the two vectors are to be considered global parameters, if one wants to interpret Eq. (31) as limiting the “free will” of the observers who choose  $\mathbf{a}, \mathbf{b}$  but preserving locality. Assuming that the unconditional probabilities are uniformly distributed, by using Bayes’s theorem one finds the conditional probability of having setting  $\mathbf{a}, \mathbf{b}$  for given  $\mathbf{u}$ ,

$$\Pi(\mathbf{a}, \mathbf{b} | \mathbf{u}) = \frac{1 - f(\mathbf{u}, -\mathbf{u}, \mathbf{a}, \mathbf{b})}{32\pi \arccos f(\mathbf{u}, -\mathbf{u}, \mathbf{a}, \mathbf{b})}. \quad (35)$$

Thus not only the settings are not free, but they are also correlated. In Appendix 10, we show how to realize this probability distribution by using two additional global unit vectors  $\mathbf{z}_A, \mathbf{z}_B$ , in terms of which the distribution is

$$\mu'(\mathbf{u}, \mathbf{v}, \mathbf{z}_A, \mathbf{z}_B | \mathbf{a}, \mathbf{b}) = \delta(\mathbf{u} + \mathbf{v}) \delta(\mathbf{a} - \mathbf{z}_A) \delta(\mathbf{b} - \mathbf{z}_B) \frac{1 - f(\mathbf{u}, \mathbf{v}, \mathbf{z}_A, \mathbf{z}_B)}{8 \arccos f(\mathbf{u}, \mathbf{v}, \mathbf{z}_A, \mathbf{z}_B)}. \quad (36)$$

Then if  $\mathbf{u}, \mathbf{v}$  are interpreted as local parameters the marginal distribution becomes

$$\mu^A(\mathbf{u} | \mathbf{z}_A, \mathbf{z}_B, \mathbf{a}, \mathbf{b}) = \frac{1 - f(\mathbf{u}, -\mathbf{u}, \mathbf{z}_A, \mathbf{z}_B)}{8 \arccos f(\mathbf{u}, -\mathbf{u}, \mathbf{z}_A, \mathbf{z}_B)}, \quad (37)$$

for  $\mathbf{a} = \mathbf{z}_A$ ,  $\mathbf{b} = \mathbf{z}_B$ , and is undeterminate in the other cases. The hypothesis of locality is satisfied by this trivial introduction of additional global parameters. We notice that it is important to distinguish the variables from the values they take. The  $\delta$  function in Eq. (36) force the value of  $\mathbf{a}$  to equal the value of  $\mathbf{z}_A$ , etc., but  $\mathbf{a}$  and  $\mathbf{z}_A$  are distinct variables.

## 7 A simple model reproducing the quantum mechanical probabilities for the spin singlet

Here we present a simple model reproducing the quantum mechanical probability for a spin-singlet. The model satisfies at the same time the hypotheses of Setting Independence, Reducibility of Correlations, and Compliance with Malus's law, while it violates the hypothesis of Uncorrelated Choice. This model was derived already in 2007, but it remained unpublished until recently,<sup>14</sup> when I came to realize that despite its simplicity it has several appealing features. In the basic formulation, the hidden-variables consist in two unit vectors  $\mathbf{u}, \mathbf{v}$ , localized at the left- and right-going particle, respectively (see Fig. 1). As one would expect for a state with zero total spin, the vectors are perfectly anti-correlated  $\mathbf{u} = \mathbf{u}_0$ ,  $\mathbf{v} = -\mathbf{u}_0$  (an index is used for the values, while the random variables are indicated by letters with no index). Furthermore, the distribution of  $\mathbf{u}$  and  $\mathbf{v}$  is tied to the polarization settings  $\mathbf{a}, \mathbf{b}$  according to

$$\mu_{\mathbf{u}=\mathbf{u}_0, \mathbf{v}=\mathbf{v}_0}(\mathbf{a}, \mathbf{b}) = \frac{1}{4} \delta(\mathbf{u}_0 + \mathbf{v}_0) \times [\delta(\mathbf{a} - \mathbf{u}_0) + \delta(\mathbf{a} + \mathbf{u}_0) + \delta(\mathbf{b} - \mathbf{v}_0) + \delta(\mathbf{b} + \mathbf{v}_0)]. \quad (38)$$

The marginal probability of a particle with  $\mathbf{u} = \mathbf{u}_0$  giving an output  $\sigma$  when impinging on a detector with polarization  $\mathbf{a}$  is given by Malus's law (which is local: no action-at-a-distance is needed)

$$M_\sigma^A(\mathbf{u}, \mathbf{a}) = \frac{1}{2} [1 + \sigma \mathbf{u} \cdot \mathbf{a}]. \quad (39)$$

At the hidden variable level, the particles are uncorrelated, i.e. the two vectors  $\mathbf{u}, \mathbf{v}$  specify all needed information, so that the conditional probability coincides with the marginal one,

$$Q_\tau^B(\mathbf{v}, \mathbf{b}, \mathbf{a}, \sigma) = M_\tau^B(\mathbf{v}, \mathbf{b}) = \frac{1}{2} [1 + \tau \mathbf{v} \cdot \mathbf{b}]. \quad (40)$$

After integration over  $\mathbf{u}_0, \mathbf{v}_0$  the quantum mechanical probability for the singlet is recovered

$$P_{\sigma, \tau}(\mathbf{a}, \mathbf{b}) = \frac{1}{4} [1 - \sigma \tau \mathbf{a} \cdot \mathbf{b}]. \quad (41)$$

Equations (39) and (40) are classical, i.e., they can be realized through classical resources. By demonstrating that Eq. (38) can be realized through classical

resources as well, the main result of this paper will be achieved: so-called quantum correlations are not exclusively quantum, since they can be mimicked by classical means. We remark that the interpretation of Eq. (38), or of any mathematical expression for probabilities, relies on a metaphysical assumption, the one about cause and effect. As will be discussed in Section 9, by interchanging cause and effect what was attributed to lack of “free will” is attributed to non-local action-at-a-distance, and vice versa.

### 7.1 Local realization through one bit of shared randomness between the measuring stations

Let us discuss how to realize Eq. (38) assuming  $\mathbf{u}, \mathbf{v}$  to be the cause, and  $\mathbf{a}, \mathbf{b}$  the effect. The two parties  $A$  and  $B$  possess identical pseudo-random\* number generators, each producing the same binary number  $c \in \{0, 1\}$ . Furthermore, by making a local measurement, they can determine, respectively,  $\mathbf{u}$  and  $\mathbf{v}$  (see Fig. 1) without perturbing the system. When  $c = 0$ ,  $A$  will measure  $\mathbf{u}$ , then flip a coin, and according to the outcome  $d$  of the flip (or at whim), will orient the polarizer either in direction  $\mathbf{u}_0$  or  $-\mathbf{u}_0$ , obtaining the outcome  $\sigma = +1$  in the first case,  $\sigma = -1$  in the second case.  $B$ , on the other hand, shall choose  $\mathbf{b}$  freely (for instance at random) when  $c = 0$ . In the other case  $c = 1$ ,  $A$  and  $B$  will reverse their actions, i.e.  $A$  shall choose  $\mathbf{a}$  freely, while  $B$  chooses  $\mathbf{b} = \mathbf{v}_0$  or  $\mathbf{b} = -\mathbf{v}_0$  according to the flipping of a coin (or at whim). Given that the conservation law  $\mathbf{u} + \mathbf{v} = 0$  holds, Eq. (38) is obtained after summing over  $c$  and  $d$ , which do not appear in the conditional probability  $P_{\sigma, \tau}(\mathbf{u}, \mathbf{v}, \mathbf{a}, \mathbf{b})$  the distribution that we have just described

$$\mu_{\lambda_0}(\mathbf{a}, \mathbf{b}) = \frac{1}{8} \delta(\mathbf{u}_0 + \mathbf{v}_0) \delta\left(\bar{c}_0(\mathbf{a} - d_0^{(A)}\mathbf{u}_0) + c_0(\mathbf{b} - d_0^{(B)}\mathbf{v}_0)\right), \quad (42)$$

with  $c \in \{0, 1\}$ ,  $\bar{c} = 1 - c$ , and  $d \in \{Heads, Tails\} = \{-1, 1\}$ .<sup>†</sup> Some authors may refer to the bit  $c$  as “non-local information”, but this is evidently just a verbal trick, since no instantaneous communication is going on, nor there is any action-at-a-distance.

It is interesting to compare two different proposed measures of “free will” in relation to this model. According to Ref. 13, putting  $\lambda = \{\mathbf{u}, \mathbf{v}, c, d\}$

$$M = \sup_{c, d} \sum \int d\mathbf{u} d\mathbf{v} |\mu_{\lambda}(\mathbf{a}, \mathbf{b}) - \mu_{\lambda}(\mathbf{a}', \mathbf{b}')| = 2. \quad (43)$$

This is the maximum value of  $M$  and corresponds to the absence of “free will”. However, since in our model half of the times  $A$  or  $B$  is free to choose the setting, while  $B$  or  $A$ , respectively, has still a binary choice,  $M$  does

---

\*By making use of classical resources, the only known way to produce two identical numbers at spacelike separated locations is to rely on a pseudo-random algorithm using an identical seed.

<sup>†</sup>Eq. (38) may also be realized in an asymmetric way, e.g., having  $A$  to always orient the polarizer along  $\pm\mathbf{u}$ .

not quantify sufficiently well the concept of “free will”. The mutual information proposed in Ref. 36, gives instead  $I(\mathbf{a}, \mathbf{b} : \mathbf{u}, \mathbf{v}, c, d) = \infty$ . However, if we discretize the allowed polarizations, having them take  $N$  possible values, we get  $I(\mathbf{a}, \mathbf{b} : \mathbf{u}, \mathbf{v}, c, d) = \log_2(N)$ , half of the maximum value  $I_{max}(\mathbf{a}, \mathbf{b} : \mathbf{u}, \mathbf{v}, c, d) = 2 \log_2(N)$ , which represents total absence of free will. This measure is hence a better quantifier of “free will”.

## 7.2 Local realization exploiting the detection loophole.

As pointed in Ref. 36, there is a close relation between the violation of Uncorrelated Choice and the detection loophole (which consists in the fact that not all pairs of particles impinging on the detectors successfully give rise to coincidence counts). This was explored in Ref. 57 to violate a Bell inequality (with only two possible settings for each detector). Indeed, an alternative explanation of the probability of the former subsection is that each particle carries with it a bit  $c_A$  (resp.,  $c_B$ ) and a unit vector  $\mathbf{u}$  (resp.,  $\mathbf{v}$ ), where the variables are anti-correlated  $c_A + c_B = 1$  (modulo 2),  $\mathbf{u} + \mathbf{v} = 0$ . The particles are programmed to act according to the following set of instructions: If  $c_A = 0$ , particle  $A$  shall behave according to Malus’s law, giving an output  $\sigma$  with the probability given in Eq. (17a). If  $c_A = 1$ , particle  $A$  will give no output unless the direction  $\mathbf{a}$  coincides (let us say within some tolerance, so that we can avoid continuous distributions) with  $\pm \mathbf{u}$ . Particle  $B$  acts in a corresponding way. If one is interested only in violating a Bell inequality, one needs two settings for  $\mathbf{a}$  and two for  $\mathbf{b}$ , so that one may limit  $\mathbf{u}$  to take eight possible values, and obtain this way an efficiency of 25% (i.e. 75% of times only one detector will give an output). Also, if one makes the model asymmetric, having, e.g., always  $c_A = 0$ , then the detection efficiency drops only to 50%, as in Ref. 57. However, if  $\mathbf{a}, \mathbf{b}$  have an uncertainty  $\Delta\Omega$  and they are spanned throughout the unit sphere, the protocol illustrated here reduces the efficiency to  $\Delta\Omega/2\pi$ .

## 7.3 Local realization of the model through shared randomness between each station and the entangler

There is no need to have shared randomness between  $A$  and  $B$ , but only between each station and the entangler. This is illustrated at length in Appendix 10.

## 7.4 Non-local realization through action-at-a-distance

Let us show, now, that inverting the cause-effect relationship between  $\mathbf{a}, \mathbf{b}$  and  $\lambda$ , the model proposed can be explained alternatively by non-local action at a distance. Indeed, Eq. (38), This could be interpreted as having, e.g.,  $\mathbf{a}$  to force the value  $\mathbf{u}_0 = \mathbf{a}$  and  $\mathbf{v}_0 = -\mathbf{a}$  when  $c_0 = 0$  and  $d_0^{(A)} = 1$ . Since  $\mathbf{v}$  is assumed to be a local parameter describing the particle at station  $B$ , we have an instance of action-at-a-distance: the setting at one station influences the parameter at the remote station.



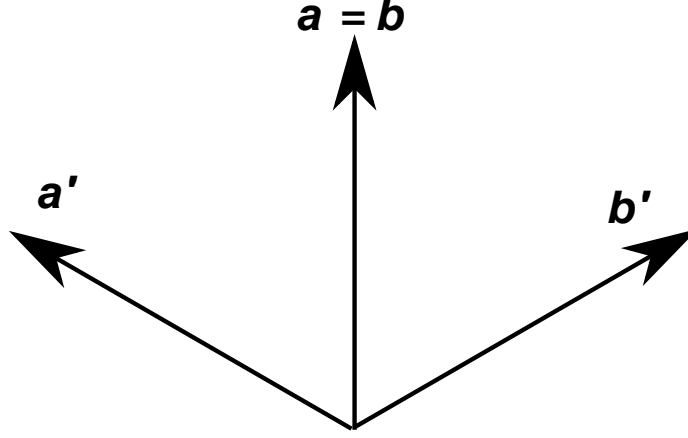


Figure 2: A possible configuration allowing  $\mathcal{E} = 4$  for the probability of Eq. (44).

## 8 A model violating Cirel'son bound.

We show that it is possible to violate the Cirel'son bound, which establishes the maximum violation of Bell inequality that quantum mechanics can achieve ( $2\sqrt{2}$ , when the Bell limit is 2). All we need to do is to mix the two models of Refs. 13 and 14: The distribution of the hidden variables is given by Eq. (38) and the  $\lambda$ -conditioned probability is not given by Malus's law, but by Eq. (33). This gives the joint probability

$$P_{\sigma,\tau}(\mathbf{a}, \mathbf{b}) = \frac{1}{4} [1 - \sigma\tau \text{sgn}(\mathbf{a} \cdot \mathbf{b})]. \quad (44)$$

The correlator is thus

$$C(\mathbf{a}, \mathbf{b}) = -\text{sgn}(\mathbf{a} \cdot \mathbf{b}), \quad (45)$$

and the Clauser-Horne parameter  $\mathcal{E} = |C(\mathbf{a}, \mathbf{b}) + C(\mathbf{a}', \mathbf{b}) + C(\mathbf{a}, \mathbf{b}') - C(\mathbf{a}', \mathbf{b}')|$  reaches the value 4 for infinitely many choices of the orientations, one of which is given in Fig. 2. This model can be realized with local classical resources along the lines of the previous section.

## 9 How to discriminate among “slave will”, “non-locality”, and “conspiracy”

The alternative realizations of the same model presented in Section 7 show that a probabilistic theory has several possible interpretations depending on the metatheoretical assumptions about the nature of its objects and events. Let us see

how one could decide in favor of one interpretation over the other. On one hand, the realization through the detection loophole can be falsified by making sure the detector's efficiency exceeds 50%. The realization through limited free will may be discerned from the one through action-at-a-distance in the following way: Let us assume that  $A$  and  $B$  can measure the parameters  $\mathbf{u}, \mathbf{v}$  (if the parameters are not experimentally accessible or controllable, there is no point in assuming they exist) during the flight of the particles, before detecting  $\sigma, \tau$ . After several runs,  $A$  will realize that half of the times  $\mathbf{u}$  is directed along  $\mathbf{a}$  or  $-\mathbf{a}$ , and that if she changes  $\mathbf{a}$ ,  $\mathbf{u}$  will change accordingly. Since the conservation law guarantees that  $\mathbf{v} = -\mathbf{u}$ ,  $A$ , assuming she has free will, will conclude that she is acting instantaneously on the remote parameter. Then she will agree with  $B$  the following instantaneous signaling protocol: they both limit their choices to four possible directions,  $\pm\mathbf{a}, \pm\mathbf{b}$ , e.g., orthogonal to each other.  $A$  fixes her setting to  $\mathbf{a}$  and  $B$  to  $\mathbf{b}$ , then  $A$  measures  $\mathbf{u}$  and  $B$  measures  $\mathbf{v}$ . If  $A$  measures  $\mathbf{u} = \mathbf{a}$ , she will quickly switch her apparatus either to  $\mathbf{b}$ , making sure that  $\mathbf{v} = -\mathbf{b}$  so that  $B$  will measure with certainty  $\tau = -1$  (which they agree to correspond to the bit 0), or to  $-\mathbf{b}$ , making sure that  $\mathbf{v} = \mathbf{b}$  so that  $B$  will measure with certainty  $\tau = +1$  (which they agree to correspond to the bit 1). The half of the times when  $\mathbf{u} = \pm\mathbf{b}$ ,  $A$  will do nothing ( $B$  knows that this is a fail, because he will see  $\mathbf{v} = \mp\mathbf{b}$  before measuring  $\tau$ ). This way,  $A$  can send to  $B$  an arbitrary sequence of bits. Now, if  $A$  is actually acting with free will and there is an action-at-a-distance, then this sequence of bits will correspond to a significant message, and will have a low entropy. If instead  $A$  has no free will, but it is the value of  $\mathbf{u}$  that determined the setting  $\mathbf{a}$ , this sequence of bits will be totally random. Of course there is always the possibility of superdeterminism, where the same mechanism that determines  $\mathbf{a}, \mathbf{b}$  is also responsible for the thousand of years of evolution of the languages, and of the recent conventional encoding of their symbols in sequences of bits. This hypothesis is clearly unfalsifiable, i.e. we cannot set its probability of being true to exactly zero, but I believe that given our experience of the world and our scientific knowledge its probability must be extremely low.

We notice that the slave will model cannot be tested unless one is able to measure or fix the additional parameters. For instance, Ref. 58 tests the free will hypothesis, but only for stochastic theories, and concedes that what is called therein “deterministic local realism” cannot be falsified. Indeed, it may happen that the random number generators and the hidden variables are correlated, but that any correlations and regularities disappear upon integrating out the hidden parameters. This would give the illusion of having two independent random number generators at the quantum level, while the randomness is lifted, or limited, at the subquantum level. Thus, for a given model that violates Uncorrelated Choice, discriminating between slave will and non-locality requires the ability to either control or detect the hidden parameters.

Finally, the realization of Section 7.1 can be interpreted indifferently as stemming from a conspiracy or from a limitation of free will. Which is the case can be decided by making sure that neither  $A$  nor  $B$  can decide  $\mathbf{a}, \mathbf{b}$  based on a measurement of  $\mathbf{u}, \mathbf{v}$ . For instance, one could ask  $A$  and  $B$  to provide before-

hand a list of the settings they will be using in each repetition. The procedure is opposite to the one followed to close the locality loophole: the settings are decided well in advance of the production of the entangled pairs, and they are kept fixed. If one observes  $A$  and  $B$  switching their apparatuses to settings not agreed upon, but coinciding with  $\pm \mathbf{u}$ , then one can conclude that either their free will is limited or that the previous decision of the sequence  $\mathbf{a}, \mathbf{b}$  influenced the future distribution of  $\mathbf{u}, \mathbf{v}$  at the entangler. At the same time, the latter hypothesis can be ruled out by having a third party to switch their apparatuses at the last moment.

## 10 Conclusions and Perspectives

Quantum correlations can be obtained by local classical resources, provided the choice of setting is correlated to the hidden variables through shared randomness or alleged action-at-a-distance. If an external observer has no access to the hidden variables, he or she can be induced to believe that genuine quantum correlations are being observed. Classical resources, however, can only provide shared pseudo-randomness, so that an infinitely intelligent observer may, in principle, discover the underlying algorithm, even without access to the hidden variables. \* On the other hand, if the external observer can measure the latter ones, he or she can discriminate if the correlations between the hidden parameters and the settings are due to a conspiracy, a lack of free will, or an action-at-a-distance.

Dropping the hypothesis of Uncorrelated Choice, however, is not the only way to reproduce the quantum mechanical correlations. One may relax the hypothesis of Setting Independence, instead, as in Ref. 12, where the correlations are simulated by having one party to transmit an information  $c$  to the other one. We have shown that this model has a natural corresponding model satisfying Setting Independence and violating Uncorrelated Choice, if one sees  $c$  as an additional hidden variable.

A third possibility consists in relaxing the hypothesis of Reducibility of Correlations.<sup>40</sup> I believe this last alternative leads to models as counter-intuitive as quantum mechanics, since they violate the classical reductionist assumption that the whole is the sum of its parts and that correlations are an expression of ignorance.

## Acknowledgements

This work was performed as part of the Brazilian Instituto Nacional de Ciência e Tecnologia para a Informação Quântica (INCT-IQ) and it was supported by

---

\*Any pseudo-random algorithm will repeat after a sufficiently large time, possibly several orders of magnitude the age of the universe even for a reasonably fast rate. However, since we are discussing a matter of principle, we may say that an infinitely patient and eternal observer, even though not necessarily intelligent, would eventually find out that the sequence is not actually random.

Fundação de Amparo à Pesquisa do Estado de Minas Gerais through Process No. APQ-02804-10.

Definitions of locality It is commonly encountered the statement that the violation of Bell inequalities implies that reality is non-local. As far as I know, locality can have any of the following meanings

- a. Impossibility of instantaneous action-at-a-distance (no action-at-a-distance, NAD).
- b. Impossibility of instantaneous communication between two space-like separated parties (no-signaling, NS).
- c. Impossibility for a body to have an arbitrary velocity (no faster-than-light speed, NFL).

These three formulations are non-equivalent, and one should specify which one is intended when using the term “locality”. For instance, the violation of (c) certainly implies the violation of (a) and (b), but the *vice versa* is not true; the violation of (b) implies the violation of (a), e.g., one could send a set of instructions to a far-away robot, that would execute some action accordingly; I surmise that it is possible to violate (a), without violating (b), even though I cannot provide presently a satisfactory example. Hence, the chain of implications holds:  $\text{NAD} \implies \text{NS} \implies \text{NFL}$ . We shall use the stricter definition of locality, NAD. We notice that there is a problem in defining what is an “action” for a stochastic theory, thus we propose a suitable extension of this concept. First, we need to distinguish between local parameters, that Bell termed “beables”, and global ones. Local parameters are attached to points in space-time, and give rise to events when they are measured in a single shot. Global parameters describe a preparation procedure, and cannot be associated with the particles composing the system. A measurement gives a collection of events  $\{e_j\}$  each corresponding to a different point of space-time. We say that a model is local (NAD) if the marginal probability of observing an event  $e_A$  at a space-time point  $A$  depends on global variables and only on those local parameters at the space-time point  $(x_A, t_A)$ .

There is a fourth definition of non-locality referring specifically to a theory predicting events  $e_j$  observed by local detectors  $\Sigma_j$  at space-like separated regions  $R_j$  given that a system is specified by some parameters  $\lambda$ ,

- d. A theory is local if the probability of observing the events  $e_j$  factorizes as

$$P(\{e\}|\lambda, \Sigma) = \prod_j P(e_j|\lambda, \Sigma_j) \quad (46)$$

In the special case of a bipartite entangled two-level system, we have that  $\Sigma_j$  represent (pseudo)spin components, and  $e_j = \pm 1$  in appropriate units. We want to show that if one takes Eq. (46) as the definition of locality then models which are patently local according to any of the definitions (a)-(c) are classified

as non-local according to (d). Let us first rewrite the left hand side of Eq. (46) by applying repeatedly Bayes's theorem

$$P(\{e\}|\lambda, \Sigma) = \prod_j P(e_j|E_j, \lambda, \Sigma), \quad (47)$$

where  $E_1 = \emptyset$  is the empty set, and  $E_j = \{e_1, e_2, \dots, e_{j-1}\}$ . Eq. (47) is a mathematical identity following from the rules of probability theory, and does not involve any physical assumption. Now, let us consider the first factor in the right hand side of Eq. (47),  $P(e_1|\lambda, \Sigma)$ . If we invoke locality (NAD) as defined above, we have the physical equality  $P(e_1|\lambda, \Sigma) = P(e_1|\lambda_G, \lambda_1, \Sigma_1)$ . The second factor, however, is  $P(e_2|e_1, \lambda, \Sigma)$ . Invoking again (NAD), we have that the marginal probability is  $P(e_2|\lambda, \Sigma) = P(e_2|\lambda_G, \lambda_2, \Sigma_2)$ . However, nothing can be said about the conditional probability  $P(e_2|e_1, \lambda, \Sigma)$ . Analogously, we have in general that

$$P(e_j|E_j, \lambda, \Sigma) = P(e_j|E_j, \lambda, \Sigma_1, \dots, \Sigma_j). \quad (48)$$

Thus, by applying (NAD) we cannot justify Eq. (46). In other words, condition (d) is stronger than (NAD). This was pointed out by Jarrett.<sup>19</sup> In order to obtain the equality in Eq. (46) we have to make a further hypothesis, which Jarrett referred to as completeness,<sup>19</sup> and Shimony as outcome-independence.<sup>20</sup> This hypothesis is simply that

$$P(e_j|E_j, \lambda, \Sigma_1, \dots, \Sigma_j) = P(e_j|\lambda, \Sigma_1, \dots, \Sigma_j), \quad (49)$$

i.e., the conditional probability of observing  $e_j$  given that  $e_1, \dots, e_{j-1}$  were observed is identical to the marginal probability of observing  $e_j$ . Bell<sup>15</sup> tried to argue that a seemingly plausible definition of “local causality” implied Eq. (49), but his proposal was criticized.<sup>17</sup>

Now, while in order to establish  $P(e_j|E_j, \lambda, \Sigma_1, \dots, \Sigma_j)$  it is necessary that observers in  $R_1, \dots, R_{j-1}$  communicate their results to the observer in  $R_j$ , the marginal  $P(e_j|\lambda, \Sigma_1, \dots, \Sigma_j)$  can be determined by means of a local measurement in  $R_j$ , without need for communication. Then, assuming that Outcome Independence holds, it is now possible to invoke once again (NAD) and obtain finally

$$P(e_j|E_j, \lambda, \Sigma_1, \dots, \Sigma_j) = P(e_j|\lambda, \Sigma_j), \quad (50)$$

which, upon substitution in the right hand side of Eq. (47) yields Eq. (46). So far, we have basically reformulated the conclusions of Ref. 19, that, however, is not widely known to physicists.

Let us establish sufficient conditions for the validity of Outcome Independence as formulated in Eq. (9). Determinism is a sufficient condition: if the knowledge of  $\lambda$  determines the outcome of the measurements  $\Sigma_j$ , all other information is redundant. Another sufficient hypothesis is that of probabilistic determinism, which is not an oxymoron, but means that knowledge of the  $\lambda$  determines the probability of the outcome, not the outcome itself. Perhaps this concept coincides with what Jarrett<sup>19</sup> calls “completeness”, which we avoid since it is a term laden with subjective meanings. Another sufficient

condition is separability:<sup>21</sup> the parameters  $\lambda = \bigcup_j \lambda_j$ , where  $\lambda_j$  are parameters attached to the particle number  $j$  and represent prepossessed values. In other words, there are no global parameters. Then we have that the probability of any outcome  $e_j$  can be determined by knowledge of the  $\lambda_j$  alone, i.e.,  $P(e_j|E_j, \lambda, \Sigma_1, \dots, \Sigma_j) = P(e_j|\lambda_j, \Sigma_j)$ , which is a special form of Outcome Independence. Bell, in particular, assumed<sup>15</sup> that all hidden parameters were of the local type (the “beables”), and that global parameters were an expression of ignorance of “beables”, and could thus be eliminated. Finally, it can be proved<sup>59</sup> that if a model satisfies factorability then there is a natural extension of this model which is deterministic.

In order to show the absurdity of identifying Eq. (46) with locality, we shall provide now a simple model which is manifestly local, but would be classified as non-local according to (d). Let us consider the following experiment: a dealer prepares two decks of cards which can be a King ( $K$ ) or a Queen ( $Q$ ), and Black ( $B$ ) or Red ( $R$ ); each deck is subdivided into pairs, such that each pair is formed by a King and a Queen, and a Black and a Red card. In the first deck ( $D_1$ ), 30% of the pairs are  $(KR, QB)$ , and 70%  $(KB, QR)$ . In the second deck ( $D_2$ ), 70% of the pairs are  $(KR, QB)$ , and 30%  $(KB, QR)$ . The dealer chooses a deck at random, with equal probability, then extracts a pair out of the deck, and handles one card each to two observers, one,  $A$ , sitting to his left and the other,  $B$ , to his right. In this model, the hidden variable  $\lambda$  is the deck which has been chosen. Once the pair is extracted from the deck,  $\lambda$  is not localized on either card, but is a global parameter, which cannot be reconstructed by observers  $A$  and  $B$  by making local observations (even if they compare their results). The only way to determine  $\lambda$  would be to check which deck was chosen at the location of the preparation. Now let us consider the joint probability that  $A$  will receive a King and  $B$  will receive a black card given that deck 1 was chosen:

$$P(K, B|\lambda = D_1) = \frac{3}{20}, \quad (51)$$

since a pair with a red King and a black Queen will be extracted with probability  $3/10$ , and the black Queen is received by  $R$  half of the times. On the other hand, if factorability holds, we should have

$$P(K, B|D_1) = P(K|D_1)P(B|D_1) = \frac{1}{4}. \quad (52)$$

Thus, according to definition (d), the model we have presented is non-local.

Notice that, elaborating further our example (or with any example which we can make out with our classical imagination), it is not possible to violate Bell inequality. The reason lies in the fact that in any classical case the hypothesis of counter-factual definiteness is always satisfied. The peculiarity of reality revealed by the experimental violation of Bell inequalities consists in the fact that no such complete characterization of a system exists in principle.

Realization of the model of Section 7 through entangler-detectors shared randomness Here we present a further realization of the model of Section 7, without having any shared randomness between  $A$  and  $B$ , but only between

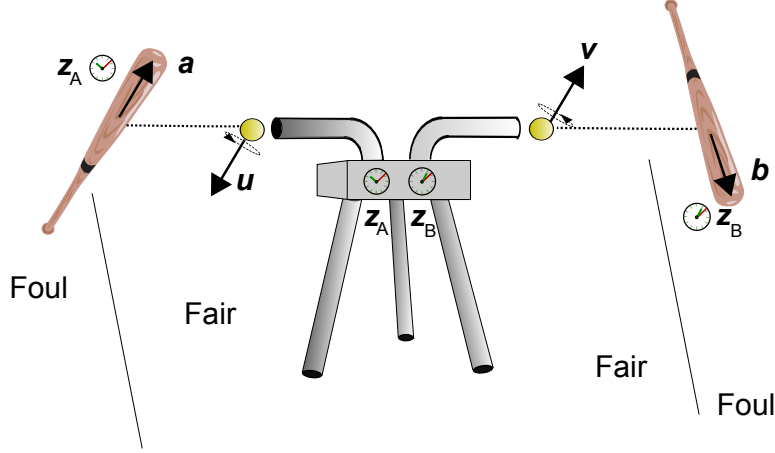


Figure 3: Scheme of the setup.

each station and the entangler. Let us consider the following setup depicted in Fig. 3: there is a sophisticated baseball pitching machine able to pitch two balls spinning with angular velocities  $\omega$  and  $-\omega$  (with  $\omega$  fixed) in opposite directions. The balls are approximatedly a rigid body and have spherical symmetry. They are pitched with fixed center-of-mass velocities in a vacuum (so that the spin does not influence the center-of-mass trajectory through the Magnus effect) such that their centers of mass follow given trajectories each ending in a bat controlled by an independent system. Each bat is carefully crafted to be a solid of revolution, its center of mass is being held fixed, and a machine varies the orientation  $\mathbf{n}$  of the bat's symmetry axis. It is empirically found that a ball pitched with  $\omega = \omega \mathbf{u}$ , after hitting a bat whose axis is oriented along  $\mathbf{n}$ , will fall in the foul ground with a frequency  $(1 - \mathbf{n} \cdot \mathbf{u})/2$ , and in the fair ground with the complementary frequency  $(1 + \mathbf{n} \cdot \mathbf{u})/2$ . Due to our politically incorrect bias against negative numbers, we shall associate the value  $\sigma = -1$  to the event of the ball being batted in the foul ground, and  $\sigma = +1$  to the alternative event. The pitching machine varies the spin of the two balls according to the following algorithm: it flips a fair coin, and according to the result, heads  $H = A$  or tails  $T = B$ , it consults either of two internal watches,  $W_A$  and  $W_B$ . Each watch is built so that the small hand has a period  $\tau_{j,s}$  and the large hand  $\tau_{j,l}$ , with  $j \in \{A, B\}$ . Unlike what happens in ordinary watches, any two of the four periods are mutually incommensurable. The positions of the hands are used to determine a unit vector  $\mathbf{z}_j$ . This may happen in a straightforward fashion, as depicted in Fig. 4 or by using a Montecarlo algorithm, so that an external observer could not predict what spin is chosen, even knowing the periods of the watches. The machine will then pitch a ball spinning along  $\mathbf{u} = \mathbf{z}_j$  to the left,

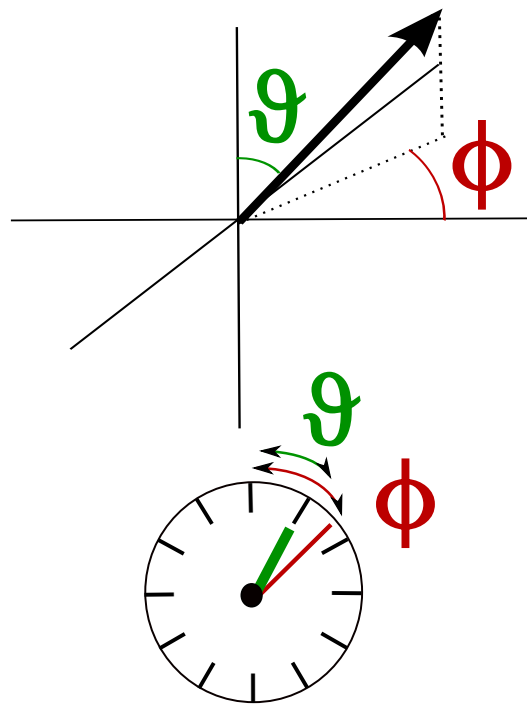


Figure 4: A unit vector is associated with the hands of a watch.



and one spinning along  $\mathbf{v} = -\mathbf{z}_j$  to the right, or *viceversa*, according to another coin flip, with outcome  $d \in \{H, T\} = \{1, -1\}$ .

The batting machine to the right possesses a watch  $W'_H$ , and the one to the left a watch  $W'_T$ , which have the same periods as the watches of the pitching machine, and are synchronized with them. Each batting machine is using the same algorithm as the pitching machine to determine the value of  $\mathbf{z}_j$ , but it is subtracting from the input values  $t_{j,h}$  the number  $\Delta t \pmod{\tau_{j,h}}$ , where  $\Delta t$  is the time-of-flight of the ball. The batting machines will orient the bats along  $\mathbf{a} = \mathbf{z}_A$  and  $\mathbf{b} = \mathbf{z}_B$ , respectively. We remind that the probability of the outcome  $\sigma$  for a ball spinning in the direction  $\mathbf{u}$  and hitting a bat oriented along  $\mathbf{n}$  is

$$M_\sigma(\mathbf{u}, \mathbf{n}) = \frac{1}{2} [1 + \sigma \mathbf{n} \cdot \mathbf{u}]. \quad (53)$$

It is easy to check that the expected joint frequency of observing the outcomes  $\sigma, \tau$  to the left and to the right when the respective bats have orientations  $\mathbf{a}, \mathbf{b}$  is

$$\begin{aligned} P_{\sigma, \tau}(\mathbf{a}, \mathbf{b}) &= \int d\mathbf{u} d\mathbf{v} \mu(\mathbf{u}, \mathbf{v} | \mathbf{a}, \mathbf{b}) M_\sigma(\mathbf{u}, \mathbf{a}) M_\tau(\mathbf{v}, \mathbf{b}) \\ &= \frac{1}{4} [1 - \sigma \tau \mathbf{a} \cdot \mathbf{b}], \end{aligned} \quad (54)$$

with the density

$$\mu(\mathbf{u}, \mathbf{v}, \mathbf{z}_A, \mathbf{z}_B | \mathbf{a}, \mathbf{b}) = \frac{1}{4} \delta(\mathbf{u} + \mathbf{v}) \delta(\mathbf{z}_A - \mathbf{a}) \delta(\mathbf{z}_B - \mathbf{b}) \sum_{j,d} \delta(\mathbf{u} - d \mathbf{z}_j) \quad (55)$$

We notice that the batting machines are devoid of “free will”, since the directions  $\mathbf{a}, \mathbf{b}$  are determined, while the pitching machine has two binary choices, first in choosing either watch, then in choosing which ball to pitch towards, e.g., left.

It is immediate to realize that the model of Ref. 13 can be reproduced by changing the algorithm used by the pitching machine. Now the pitching machine will use both watches  $W_H, W_T$  to determine two unit vectors  $\mathbf{a}, \mathbf{b}$ . The pitcher is in possession of a third watch  $W_0$ , whose hands determine a unit vector  $\mathbf{u}$ . The third watch, however, is not ticking regularly, but it is coupled to the watches  $W_R$  and  $W_L$  in such a way that the hands correspond to the vector  $\mathbf{u}$ , for given  $\mathbf{a}, \mathbf{b}$ , with the frequency

$$\Pi(\mathbf{u} | \mathbf{a}, \mathbf{b}) = \frac{1 - f(\mathbf{u}, -\mathbf{u}, \mathbf{a}, \mathbf{b})}{8 \arccos f(\mathbf{u}, -\mathbf{u}, \mathbf{a}, \mathbf{b})}. \quad (56)$$

The pitcher will then proceed to pitch two balls, the first, spinning about  $\mathbf{u}$ , to the left, and the second, spinning about  $-\mathbf{u}$ , to the right. The other parameters of the procedure  $\omega, v$  are readjusted such that a ball spinning around  $\mathbf{u}$  hitting a bat oriented along  $\mathbf{n}$  will give the outcome  $\sigma$  with probability

$$P_\sigma(\mathbf{u}, \mathbf{n}) = \delta_{\sigma, \text{sgn}(\mathbf{u} \cdot \mathbf{n})}. \quad (57)$$

The batters, on the other hand, keep using the former algorithm in order to determine  $\mathbf{a}, \mathbf{b}$ .

## References

## References

- [1] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777–780 (1935).
- [2] J. S. Bell, *Physics* **1**, 195–200 (1964).
- [3] J. F. Clauser and M. A. Horne, *Phys. Rev. D* **10**, 526–535 (1974).
- [4] S. J. Freedman and J. F. Clauser, *Phys. Rev. Lett.* **28**, 938–941 (1972).
- [5] A. Aspect, P. Grangier, and G. Roger, *Phys. Rev. Lett.* **49**, 91–94 (1982).
- [6] A. Aspect, J. Dalibard, and G. Roger, *Phys. Rev. Lett.* **49**, 1804–1807 (1982).
- [7] P. R. Tapster, J. G. Rarity, and P. C. M. Owens, *Phys. Rev. Lett.* **73**, 1923–1926 (1994).
- [8] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, *Phys. Rev. Lett.* **81**, 5039–5043 (1998).
- [9] W. Tittel, J. Brendel, B. Gisin, T. Herzog, H. Zbinden, and N. Gisin, *Phys. Rev. A* **57**, 3229–3232 (1998).
- [10] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, *Nature* **409**, 791–794 (2001).
- [11] A. Schopenhauer. *The world as will and representation*. (1819). [E. F. J. Payne transl., Dover Edition, 1969].
- [12] B. F. Toner and D. Bacon, *Phys. Rev. Lett.* **91**, 187904 (2003).
- [13] M. J. W. Hall, *Phys. Rev. Lett.* **105**, 250404 (2010).
- [14] A. Di Lorenzo, *Journal of Physics A: Mathematical and Theoretical* **45**, 265302 (2012).
- [15] J. S. Bell, *Epist. Lett.* **9**, 11–24 (1976).
- [16] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Phys. Rev. Lett.* **23**, 880–884 (1969).
- [17] A. Shimony, M. A. Horne, and J. F. Clauser, *Epist. Lett.* **13** (1976).
- [18] J. S. Bell, in *49th International School of Physics “Enrico Fermi”: Foundations of quantum mechanics, Varenna, Italy*, ed. B. D’Espagnat, 171–181, Academic Press, New York (1971).
- [19] J. P. Jarrett, *Noûs* **18**, pp. 569–589 (1984).

- [20] A. Shimony, in *Sixty-Two Years of Uncertainty: Historical, Philosophical, and Physical Inquiries into the Foundations of Quantum Mechanics.*, ed. A. I. Miller, 33–43, Plenum, New York (1990).
- [21] D. Howard, in *Philosophical Consequences of Quantum Theory: Reflections on Bell’s Theorem*, eds. J. Cushing and E. McMullin, 224–253, University of Notre Dame Press (1989).
- [22] J. P. Jarrett, in *Philosophical Consequences of Quantum Theory: Reflections on Bell’s Theorem*, eds. J. Cushing and E. McMullin, 60–79, University of Notre Dame Press (1989).
- [23] M. L. G. Redhead, in *Philosophical Consequences of Quantum Theory: Reflections on Bell’s Theorem*, eds. J. Cushing and E. McMullin, 145–153, University of Notre Dame Press (1989).
- [24] A. Leggett, *Foundations of Physics* **33**, 1469–1493 (2003).
- [25] S. Gröblacher, T. Paterek, R. Kaltenbaek, Č. Brukner, M. Żukowski, M. Aspelmeyer, and A. Zeilinger, *Nature* **446**, 871–875 (2007).
- [26] C. Branciard, N. Brunner, N. Gisin, C. Kurtsiefer, A. Lamas-Linares, A. Ling, and V. Scarani, *Nature Physics* **4**, 681–685 (2008).
- [27] T. Paterek, A. Fedrizzi, S. Gröblacher, T. Jennewein, M. Żukowski, M. Aspelmeyer, and A. Zeilinger, *Phys. Rev. Lett.* **99**, 210406 (2007).
- [28] C. Branciard, A. Ling, N. Gisin, C. Kurtsiefer, A. Lamas-Linares, and V. Scarani, *Phys. Rev. Lett.* **99**, 210407 (2007).
- [29] M. D. Eisaman, E. A. Goldschmidt, J. Chen, J. Fan, and A. Migdall, *Phys. Rev. A* **77**, 032339 (2008).
- [30] J. Romero, J. Leach, B. Jack, S. M. Barnett, M. J. Padgett, and S. Franke-Arnold, *New Journal of Physics* **12**, 123007 (2010).
- [31] M. Paternostro and H. Jeong, *Phys. Rev. A* **81**, 032115 (2010).
- [32] C.-W. Lee, M. Paternostro, and H. Jeong, *Phys. Rev. A* **83**, 022102 (2011).
- [33] A. Aspect, *Nature* **446**, 866–867 (2007).
- [34] A. Di Lorenzo, *Reassessment of Leggett inequality*, preprint, <http://arxiv.org/abs/1104.1695v1>, (2011).
- [35] C. H. Brans, *International Journal of Theoretical Physics* **27**, 219–226 (1988).
- [36] J. Barrett and N. Gisin, *Phys. Rev. Lett.* **106**, 100406 (2011).
- [37] N. J. Cerf, N. Gisin, S. Massar, and S. Popescu, *Phys. Rev. Lett.* **94**, 220403 (2005).

- [38] N. Brunner, N. Gisin, S. Popescu, and V. Scarani, *Phys. Rev. A* **78**, 052111 (2008).
- [39] M. Pawłowski, J. Kofler, T. Paterek, M. Seevinck, and Č. Brukner, *New J. Phys.* **12**, 083051 (2010).
- [40] A. Di Lorenzo, *Determination of hidden variable models reproducing the spin-singlet*, preprint, <http://arxiv.org/abs/1105.1286v4>, (2011).
- [41] G. Boole, *Philos. Trans. R. Soc. Lond.* **152**, 225–252 (1862).
- [42] J. Bass, *C.R. Acad. Sci. Paris* **240**, 839–841 (1955).
- [43] E. D. Schell, in *Proc. 2nd Symp. Lin. Programming*, ed. H. A. Antosiewicz, number 2, 615–642, DCS/Comptroller H.Q. U.S.A.F., Washington (1955).
- [44] N. N. Vorob’ev, *Dokl Akad. Nauk SSSR* **124**, 253–256 (1959).
- [45] N. N. Vorob’ev, *Theory of Probability and its Applications* **7**, 147–163 (1962).
- [46] E. C. Kemble, *Phys. Rev.* **47**, 973–974 (1935).
- [47] A. E. Ruark, *Phys. Rev.* **48**, 466–467 (1935).
- [48] N. Bohr, *Phys. Rev.* **48**, 696–702 (1935).
- [49] H. C. Wolfe, *Phys. Rev.* **49**, 274 (1936).
- [50] W. H. Furry, *Phys. Rev.* **49**, 393–399 (1936).
- [51] W. H. Furry, *Phys. Rev.* **49**, 476 (1936).
- [52] I. Pitowsky, *Brit. J. Phil. Sc.* **45**, 95–125 (1994).
- [53] K. Hess and W. Philipp, *Foundations of Physics* **35**, 1749–1767 (2005).
- [54] A. Fine, *Phys. Rev. Lett.* **48**, 291–295 (1982).
- [55] L. Accardi and M. Regoli, in *Foundations of probability and physics. Proceedings of the conference, Växjö, Sweden, 25 Nov – 1 Dec 2010*, ed. A. Khrennikov, 1–28, World scientific (2001).
- [56] T. Durt, *Foundations of Physics* **27**, 415–434 (1997). 10.1007/BF02550164.
- [57] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, V. Scarani, V. Makarov, and C. Kurtsiefer, *Phys. Rev. Lett.* **107**, 170404 (2011).
- [58] T. Scheidl, R. Ursin, J. Kofler, S. Ramelow, X.-S. Ma, T. Herbst, L. Ratschbacher, A. Fedrizzi, N. K. Langford, T. Jennewein, and A. Zeilinger, *Proc. Natl. Acad. Sci. USA* **107**, 19708–19713 (2010).
- [59] M. J. W. Hall, *Phys. Rev. A* **84**, 022102 (2011).